

## A METHOD AND ARRANGEMENT FOR BROWSING DOCUMENTS IN A DATABASE

FIELD OF THE INVENTION

5 The present invention relates to a method of browsing documents stored in a database. The invention relates further to an arrangement for performing the browsing operations.

10 BACKGROUND OF THE INVENTION

15 There are several known techniques for storing a large number of documents in database. A single document in the database may contain references to other documents and for instance to image (graphics) files and/or sound files of the same or even to a separate database.

20 An example of the kind of document stored in such a database is an HTML document (HyperText Markup Language) ie. the "hyperdocument", which is widely used in an Internet environment. The "Internet" is a worldwide network of computers capable of communication with each other by means of a suitable protocol, such as TCP/IP (Transmission Control Protocol / Internet Protocol).

25 During recent years the use of the Internet has expanded rapidly, especially after the introduction of so called Web browsers which allow a simple graphical user interface based (GUI) access to network servers supporting HTML documents.

30 Typically, a database of HTML documents is stored on a Web server connected to the WWW (World Wide Web). The WWW can be understood in turn as a collection of servers connected to the Internet, which utilise so-called  
35 HyperText Transfer Protocol (HTTP). This HTTP in turn utilises said HTML standard page description.

It is possible to browse separate documents in given locations by a suitable terminal device, ie. a "Web browser", which can consist of a microcomputer or a workstation. When the Web browser is connected to the WWW, a Web server receives an URL (Uniform Resource Locator) request from the browser, decodes the URL, handles the document files and sends the requested files to the Web browser.

It is also possible to browse documents locally, ie. in a local file system of a data processing device which may not be connected to the WWW - i.e. a so called stand-alone data processing device. In this case, the document address (ie. a local file path) is given to the local file system which then retrieves the file for the browser. By using relative document references it is possible to transfer the whole document collection to a different location without any modifications to the documents itself.

#### SUMMARY OF THE INVENTION

The above arrangements involve some problems. Existing browsing means, eg. Web browsers, are not enabled to browse a document database which consists of a set of documents stored as a single file. Rather, the browser can only access separate documents located at a given address (URL). However, a large collection of such individual documents, such as thousands of documents, is not easily managed. A single file could be managed far more easily than a large collection of separate files, which files may even be of different types. A single file could also be more easily given a product identity and a version identifier, and the quality of the information contained in the database could be guaranteed through proper version handling.

As the documents in a single file database typically contain links to other documents in the database, a problem arises in enabling the transfer of the database to different locations while keeping the database in a manageable condition even in these different locations. In other words, a problem lies in the defining and handling of the document links so that a location independent database can be achieved. It would be a further advantage if the single file database could be defined such that the same arrangement could be used both in network use (the same database could be copied to any network server) and in a local file system, such as in a stand alone microcomputer (the same database could be copied to any file path in the file system).

Therefore, it is an object of the present invention to overcome the disadvantages of the prior art solutions and to provide a new type of solution for browsing of documents stored in a single file database.

An object of the present invention is also to provide a browsing method and arrangement by means of which it is possible to have a complete document database (eg. a document library, a document archive etc.), which is formed as a single file, which enables the browsing of this database in a network environment and/or in a stand-alone data processing device.

An object of the present invention is also to provide a method and arrangement in which the document references are defined and handled such that a location independent single file database is achieved.

It is also an object of the present invention to provide a database arrangement and method which can be utilised both in a network environment and in a stand-alone data

processing device, and in addition such that the documents are usable in both occasions. A further object is an arrangement for this in which two programs can handle such a database.

5

Other objects and advantages of the present invention will be brought out in the following part of the specification taken in conjunction with the accompanying drawings.

10

The objects are obtained by the disclosed method for browsing documents stored in a database. In said method at least one of said documents includes references to other documents in the database, and the browsing  
15 comprises steps for retrieving of a desired document from the database in response to a request for that document by browsing means, dynamically transforming the references of said retrieved document from a special syntax to a form said browsing means is capable of  
20 understanding, and transmitting the referenced documents to said browsing means.

According to an aspect of the present invention there is provided a method of browsing a set of documents stored  
25 electronically as a single file, at least one of the documents containing links to others of the documents, the method comprising:

retrieving said at least one document from the database in response to a request for that document by  
30 an electronic document browsing means;

scanning the retrieved document to identify said links;

transforming the links into a format which is recognisable by the document browsing means; and

35 transmitting the retrieved document including said transformed links to the browsing means.

According to this aspect of the invention, it is possible to provide a database as a single file which can be easily handled, i.e. stored, copied, etc. This avoids individual documents being confused, lost, or inadvertently amended. The database may be stored in a compressed form, e.g. as a ZIP file. By providing for the transformation of internal document links into, for example, URLs, these links are made independent of the actual storage location of the database. The transformation is handled by a proprietary "interface" application.

It is to be understood that the term "document" used above is intended to indicate a variety of different forms and is not limited to text documents. For example, documents may include images, video and audio sequences, and graphics.

According to an embodiment of the invention, a method is disclosed for browsing documents stored in at least two databases connected to each other by communication means, wherein at least one of said documents is including references to other documents and/or files in one or several of the databases. The browsing comprises steps for retrieving of a desired document from at least one of the databases as response to a request for that document by browsing means, a dynamic transforming of the references of said retrieved document from a special syntax to a form said browsing means is capable of understanding, and transmitting the referenced documents and/or files to said browsing means.

According to a preferred embodiment, an arrangement is provided for browsing documents stored in a database, wherein at least one of said documents to be browsed

includes references to other documents and/or files in the database, the arrangement comprises browsing means, means for retrieving a desired document from the database and for accomplishing a dynamic transformation of the references of said desired document from a special syntax to a form said browsing means is capable of understanding, and means for transmitting said referenced documents and/or files to said browsing means.

10

A further embodiment discloses a method for browsing operation and maintenance documents for telecommunication systems stored in a database, at least one of said documents including references to other documents and/or files in the database, wherein the browsing comprises retrieving of a desired document from the database as response to a request for that document by browsing means, a dynamic transforming of the references of said retrieved document from a special syntax to a form said browsing means is capable of understanding, and transmitting the referenced documents and/or files to said browsing means.

Several advantages are obtained by means of the present invention, since the solution provides a simple, reliable and controllable manner for document browsing. The solution provides an automatic and transparent linking of various different documents. The document database to be browsed is provided in a form of a complete document database. It also enables the use of one single database in different situations, such as in a network (eg. world wide Web of the Internet) and standalone (eg. PC) use, and on different physical locations without a need to be modified. This all provides an improved and eased administration and maintenance of such databases compared to those having thousands of separate documents of various types.

In the following the present invention and the other objects and advantages thereof will be described by way of an example with reference to the annexed drawings, in which similar reference characters throughout the various Figures refer to similar features. It should be understood that the following description of an example of the invention is not meant to restrict the invention to the specific forms presented in this connection but rather the present invention is meant to cover all modifications, similarities and alternatives which are included in the spirit and scope of the present invention, as defined by the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic presentation of an arrangement embodying the present invention for network use;

Figure 2 is a schematic presentation of an arrangement according to one embodiment of the present invention for stand alone use;

Figure 3 is a schematic presentation of an arrangement embodying the present invention for both the network and stand alone use;

Figure 4 discloses an exemplifying flow chart of the operations in Web use according to one embodiment of the present invention; and

Figure 5 discloses one example of a page displayed to the user.

DETAILED DESCRIPTION OF CERTAIN EMBODIMENTS

The present invention provides a new type of method and arrangement for browsing a database, either in a network environment, such as in the world wide Web (WWW) (Figure 1), or in a stand-alone data processing device (Figure 2) such as a microcomputer or personal computer (PC), or both (Figure 3). The database can also be located in LAN (Local Area Network).

Suitable data processing devices or computers are per se known by the skilled person, and thus it is only mentioned herein that such apparatus usually comprises a central processing unit (CPU), memory means such as RAM (Random Access Memory) and ROM (Read Only Memory), and a display or corresponding means for presenting information to the user. A keypad and a mouse are provided for enabling the user to control and instruct the apparatus.

Figure 1 is a schematic presentation of one embodiment according to the present invention. In Figure 1, a standard Web server (www server) is indicated by reference numeral 2. A document server 3 is arranged to handle a document database 4, and to transform document references in a manner to be described below. In Figure 1, single lined arrows indicate requests and double lined arrows indicate data communication.

A Web browser 6 is arranged to communicate with the Web server means 2. The skilled person knows the operational principles and operation of different Web browsers, such as Netscape Navigator™ or Explorer™ by Microsoft Corp., and thus they are not explained in more detail herein. An advanced Web browsing system in connection with the Internet is also disclosed by US patent 5,572,643 (Judson), which is incorporated herein by reference.



The communications between the WWW server 2 and the Web browser 6 is arranged to use http (hyper text transfer protocol) and a TCP/IP connection (Transport Control  
5 Protocol/Internet Protocol).

The document database 4 consists of a single ZIP file (i.e. with a single filename recognisable to the operating system of the document server) containing a  
10 number of HTML documents (HyperText Markup Language), ie. "hyperdocuments", which are listed in an index at the start of the database. A document may contain references to another document according to a specific syntax or format.

15 The document server 3 is provided with capability for managing the document database 4, and is arranged to dynamically transform references within the various documents from a specific syntax to a form which is  
20 understandable by the Web browser 6. One example of this kind of transforming operation is shown by a detailed exemplifying code listing included in this description.

25 In more detail, the Web (eg. World Wide Web: WWW) server 2 can be accessed through a graphic user interface, such as, eg. Mosaic™, Netscape™, Cello™ etc., for reading the HTML documents. In order to find the requested document to be browsed, the document reference or link  
30 sent from the Web browser 6 to the Web server 2 must be in a Uniform Resource Locator (URL) form. The URL has the following general syntax:

<protocol>:<address>?<data>#<name>

It is, however, to be understood that the above syntax is given only as an example, and that other possible syntaxes could also be used.

- 5 The browser 6 is provided with knowledge about a predefined set of protocols, and their respective protocol handlers. These protocols are per se known, and are such as the above mentioned 'http' (hyper text transfer protocol) and 'file'.

10

- When the 'http' protocol is used, the address will be in the form of an Internet address (eg. lm.ericsson.se) and a pathname to the requested document in the Web server 2 (eg. /docs/doc.html). The request is routed over the
- 15 Internet to a specified host machine which is positioned at the Internet address given to the URL. At the host machine, the Web server uses the pathname portion of the address to locate the requested document. The pathname may also specify a CGI (Common Gateway Interface)
- 20 program to be executed, in which case the CGI program is responsible for returning the requested document to the Web server 2, which in turn will transmit it to the Web browser means 6 (the client) through the TCP/IP connection.

25

- In the case where 'file' is used, the address is a local file path (eg. C:\mydir\doc.htm). The handler of this type of address is the local file system, which is capable of fetching the requested document on the basis
- 30 of the given file path, and giving it to the Web browser 6 for displaying purposes.

- In a specific case where a Netscape™ compatible browser is used, it is possible for an external program to
- 35 define (name) a new protocol to the browser means 6. This external program implicitly becomes the protocol handler for this new protocol. When such a protocol has

been defined, the browser sends requests (URLs) containing this protocol to the external program (eg. program A in Figure 2). The external program will then become responsible for decoding this type of URL.

5

The document references in the documents included within the single file database contain a special protocol (named as 'edw') which is not automatically known by the browser means 6. For this reason certain

10 transformations of the references (URLs) must take place before a document in the database can be located and given to the browser means 6 for display. The situation is the same both in network use (Figure 1) and stand alone use (Figure 2).

15

The following is discussed in view of these transformations. Reference is also made to the following exemplifying code listing:

20 In the database (4):-----

```
<BODY BACKGROUND="edw:/alex?AC=image&FN=bopi_bg.gif">
```

```
<A NAME=TOP></A>
```

```
<IMG WIDTH="98%" HEIGHT=71 SRC="edw:/alex?AC=image&FN=bopihead.gif">
```

25 <P><BR>

```
<P>
```

```
<A NAME=CONTENTS><B>Contents</B></A>
```

```
<PRE>
```

```
<A HREF="edw:/alex?FN=bopi_in1.html#CHAPTER1"><B>GENERAL,APT</B></A>
```

30 <A HREF="edw:/alex?FN=bopi\_in1.html#CHAPTER1.1">Administrative routines</A>

```
<A HREF="edw:/alex?FN=bopi_in1.html#CHAPTER1.2">Product handling</A>
```

In Web browsing:-----

35 <BODY BACKGROUND="http://lm.ericsson.se/cgi-bin/alexserv?DB=aa33.alx&AC=image&FN=bopi\_bg.gif">

```

<A NAME=TOP></A>
<IMG WIDTH="98%" HEIGHT=71
SRC="http://lm.ericsson.se/cgi-bin/alexserv?DB=aa33.alx&AC=imag e&FN =bopihead.gif">
<P><BR>
5  <P>
  <A NAME=CONTENTS><B>Contents</B></A>
  <PRE>
    <A
      HREF="http://lm.ericsson.se/cgi-bin/alexserv?DB=aa33.alx&FN=bopi_in1.html
10  #CHAPTER1"><B>GENERAL, APT</B></A>
    <A
      HREF="http://lm.ericsson.se/cgi-bin/alexserv?DB=aa33.alx&FN=bopi_in1.html
      #CHAPTER1.1">Administrative routines</A>
    <A
15  HREF="http://lm.ericsson.se/cgi-bin/alexserv?DB=aa33.alx&FN=bop
    i_in1.html#CHAPTER1.2">Product handling</A>

In stand alone (PC) browsing:-----

20  <BODY BACKGROUND="file:///C:/TEMP/ALEX_TMP/aa33/alex0022.gif">
  <A NAME=TOP></A>
  <IMG WIDTH="98%" HEIGHT=71 SRC="file:///C:/TEMP/ALEX_TMP/aa33/alex0023.gif">
  <P><BR>
  <P>
25  <A NAME=CONTENTS><B>Contents</B></A>
  <PRE>
    <A HREF="edw:/alex?FN=bopi_in1.html#CHAPTER1"><B>GENERAL, APT</B></A>
    <A HREF="edw:/alex?FN=bopi_in1.html#CHAPTER1.1">Administrative routines</A>
    <A HREF="edw:/alex?FN=bopi_in1.html#CHAPTER1.2">Product handling</A>
30

In view of above exemplifying listings, and when
supposed that the address to a document in the database
2 is:

35  edw:/alex?FN=<doc name>,

```

wherein <doc name> is the name of the referenced document. It is to be noted that the name of the database is not included in the address, since the user may rename the database file.

5

The address to the document that should be retrieved from a Web server has the following form:

`http://<server address>/<server name>?FN=<doc name>&DB=<db name>`

10

wherein <server address> is the Internet address of the Web server in which the database is located, <server name> is the name of the server document program and <db name> is the name of the document database. When a document is retrieved from a Web located database, all URLs (both to other documents and to image files included in with document) the document contains must be transformed to the above described form.

15

In case the document is retrieved via the stand alone program, ie. via 'A' in Figure 2, the document cannot be given directly to the browser means 6. This is a restriction of the external interface of the browser. Thus the document must be firstly stored as a file on a local storage means, such as a local hard disk (C:) 10. Therefore the address to the file retrieved in the stand alone mode has the following form:

20

`file://C:\temp\<file name>,`

30

where 'C:\temp\' can be any local file path and 'file name' is the name of the temporary file containing the actual requested document. When a certain document is retrieved from a single file database 4, all such URLs which make a reference to other documents in the database can be left as they are, since they will be decoded in due course by stand alone program 'A' at 12.

35

However, all the references the document makes to image files must be transformed to the above form, as the browser means 6 expects them to be immediately available. This is because of the external interface of the browser, which sets a restriction according to which the mechanism for requesting URLs from the external program works only for complete documents, and not for images included in a certain document. Thus the external program 'A' must retrieve all images the document contains simultaneously, as it retrieves the text file of the document.

It is to be understood that all shown means of Figure 2 are preferably disposed with a single data processing device such as a PC. The database 4 may also be a part of a LAN network (Local Area network).

Now, after having described the formations, the actions to be taken by the document browsing arrangement will be described in more detail for both the network version and the stand alone version, respectively.

The operation in network environment will be discussed first referring again to Figure 1. The operation is initiated by taking a connection to document server 3. A list of databases including the database 4 is given through the display of the browser 6, whereafter the user may select the desired database (database 4 in the example). Thereafter a supposition document from the selected database is displayed (eg. the document of Figure 5). The document server 3 receives then a request for a certain document from the browser means 6 via Web server means 2, as is indicated by the single lined arrows. The document server retrieves then the document from the database 4 using the database index.

The document is first unzipped and is then scanned by the document server to identify all references to other documents in the database. These references are then transformed from the proprietary edw references into  
5 'http' form references including the actual address to the document server and the database name. The document is then given to the Web server 2, which transmits it further to the browser means 6, as is indicated by the double lined arrows, respectively.

10

The browser means 6 loads the document and upon finding references to possible image files sends out a new request to retrieve each of the referenced image files and similar. Since the reference to a certain image  
15 file is an http link, each reference will generate a new request for the document server 3 to retrieve and unzip the image file (or any other possible referenced file).

The operation steps for the network use are also  
20 illustrated by the flow chart of Figure 4. At step 100 the browser means send a request for a certain document to the Web server. The document is then retrieved from the database by means of the document server, step 102. After unzipping the document, the references to other  
25 documents and/or files are searched, step 104. In the next step (106), the syntaxes of the found references are transformed by means of the document server into a form the browser is capable of understanding. Thereafter the document is given to the Web server and  
30 is further transmitted to the browser means as a response to the request (steps 108 and 110, respectively). The document is loaded into the browser, step 112, and possible requests for the additional referenced documents are sent to the Web server, step  
35 114. These are retrieved from the database (or several databases), unzipped, and returned in a corresponding manner to the browser, step 116.

The document is now retrieved to the browser means 6 and displayed by the display means thereof, step 118. One example of such a document is shown in Figure 5. The documents and/or files can be browsed locally at the browser. As shown in Figure 5, the document may contain some additional references shown to the user which are of the above mentioned http type, and thus when the user 'clicks' eg. by the mouse or buttons of the keypad one or several of the references, the whole process will begin again.

In the stand alone use (Figure 2), the operation can be described in the following manner. The user starts program 'A' which starts the actual program in case it is not already running. The program 'A' defines itself to the browser 6 as the handler for the 'edw' protocol and will receive all requests that are based on this protocol. So when the user clicks on a reference which is utilising the 'edw' protocol, the whole URL is sent to the program 'A'. The URL contains the name of the document to be retrieved.

The document is retrieved from the database, unzipped, and possible links to image files or the like are searched and the found file links are transformed to 'file:/' links. Simultaneously the image and other possible files are retrieved from the database and stored on the local file system in a place which corresponds to the 'file:/' links inserted into the document. The text document is given a file name and stored to the local hard disk 10. Finally the browsing means 6 is instructed to load the file from with hard disk 10.

The browsing means 6 loads the file and upon finding searched references to the appropriate files it loads



also these into the window or display of the browsing means 6. The document in the browser display contains now the document references of the 'edw' type described above, and thus, when the user clicks on one of them a  
5 new request will be generated to the external program A. Thus the stand alone data processing device is capable of handling the document database, to communicate with the browser's external interface and to transform the references to the image file documents or similar from  
10 their special syntax to a form which the browser is able to understand.

Figure 3 illustrates the possibility of using both of the above described arrangements... This could be the  
15 case, eg. with portable computers which may be connected for most of the time to the communications network, but are occasionally used as stand-alone computers.

According to the present invention, it is possible to  
20 transmit several documents and/or files from more than one database as a single unit from the Web server to the browsing means.

According to one embodiment the files are transmitted  
25 and handled as a single compressed file during the retrieving operations. The compressing may be accomplished, eg. by means of a well known ZIP program.

The invention enables the use of one single database in  
30 different situations, such as in a network and stand-alone use, and on different physical locations without a need to be modified. This provides an improved and eased administration and maintenance of such databases.

35 Thus one preferred use for the inventive solution is browsing of operation and maintenance documents for telecommunication equipment (eg. AXE O&M documents, AXE

being a trademark of Telefonaktiebolaget L M Ericsson for an exchange by the same). An advantage to the operator is that only one single document database is required to be installed in one single location in the Intranet of the operator, whereafter the documentation will become available throughout the whole organisation. On the other hand, when a service person is going to go to, eg. accomplish some repairs in an exchange at which no Web connections are available, the service person can have the latest documentation in a stand-alone portable computer he/she is carrying.

Thus, the invention provides an apparatus and a method by which a significant improvement can be achieved in the area of Web browsing. It is to be noted that the foregoing examples of the embodiments of the invention are not intended to restrict the scope of the invention defined in the appended claims. All additional embodiments, modifications and applications obvious to those skilled in the art are thus included within the spirit and scope of the invention as set forth by the claims appended hereto.